
UNIVERSITI SAINS MALAYSIA

Second Semester Examination
2015/2016 Academic Session

June 2016

CPT443 – Automata Theory & Formal Languages
[Teori Automata & Bahasa Formal]

Duration : 3 hours
[Masa : 3 jam]

INSTRUCTIONS TO CANDIDATE:

[ARAHAN KEPADA CALON:]

- Please ensure that this examination paper contains **SEVEN** questions in **NINE** printed pages before you begin the examination.

*[Sila pastikan bahawa kertas peperiksaan ini mengandungi **TUJUH** soalan di dalam **SEMBILAN** muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]*

- Answer **ALL** questions.

*[Jawab **SEMUA** soalan.]*

- You may answer the questions either in English or in Bahasa Malaysia.

[Anda dibenarkan menjawab soalan sama ada dalam bahasa Inggeris atau bahasa Malaysia.]

- In the event of any discrepancies, the English version shall be used.

[Sekiranya terdapat sebarang percanggahan pada soalan peperiksaan, versi bahasa Inggeris hendaklah diguna pakai.]

1. (a) Let L_1 be the language {laptop, hp, tablet} and L_2 be the language {battery, cable, ϵ }.

Biar L_1 diberi sebagai bahasa {laptop, hp, tablet} dan L_2 diberi sebagai bahasa {battery, cable, ϵ }

- (i) How many elements are there in L_1^* ?

Berapakah elemen yang terdapat dalam L_1^ ?*

- (ii) List five elements in L_2^+ .

Senaraikan lima elemen yang terdapat dalam L_2^+ .

- (iii) List the elements of $L_1 L_2$ in lexicographic order.

Senaraikan elemen-elemen untuk $L_1 L_2$ mengikut urutan lesikal.

(5/100)

- (b) For each of the following statements, state whether it is **True** or **False**. If False, explain why. You may use an example to illustrate the explanation.

*Untuk setiap pernyataan di bawah, nyata sama ada ia adalah **Betul** atau **Salah**. Jika Salah, terangkan sebabnya. Anda boleh menggunakan contoh untuk gambarkan penerangan anda.*

- (i) $\forall L ((L^+)^+ = L^+)$.

- (ii) $\forall L ((L^*)^+ = (L^+)^*)$.

- (iii) $\forall L (L^* = L^+ \cup \emptyset)$.

- (iv) $\forall L (\{\epsilon\} \cup L^+ = L^+)$.

- (v) $\forall L_1, L_2$ where $L_1 \neq L_2 ((L_1 L_2) \neq (L_2 L_1))$.

(5/100)

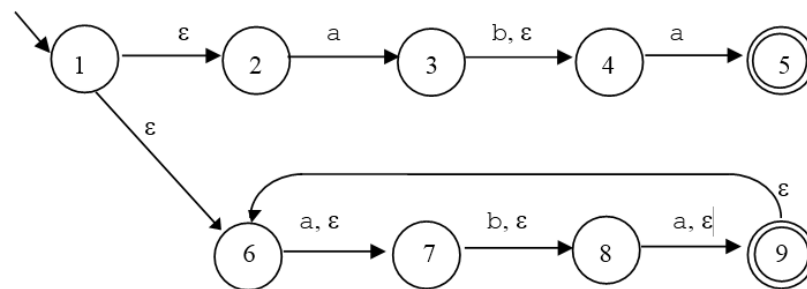
2. (a) Convert the following Nondeterministic Finite Automaton (NFA) into a Deterministic Finite Automaton (DFA) through the subset construction. Show the value of $\text{eps}(q)$ for each state q . Describe the transition function using a transition table. You may list down the functions for related states only. Present your DFA using a transition diagram.

Ubah Automaton Terhingga Tak Berketentuan (NFA) berikut kepada Automaton Terhingga Berketentuan (DFA) melalui pembinaan subset.

Tunjuk nilai $\text{eps}(q)$ untuk setiap keadaan q .

Terangkan fungsi peralihan dengan menggunakan jadual peralihan. Anda boleh menyenaraikan fungsi untuk keadaan yang berkenaan sahaja.

Bentang DFA anda dengan menggunakan gambarajah peralihan.



(6/100)

- (b) Design a Deterministic Finite Automaton (DFA) for the following languages. Present your DFA by a transition diagram.

Reka suatu Automaton Terhingga Berketentuan (DFA) untuk bahasa-bahasa berikut. Bentang DFA anda dengan menggunakan gambar rajah peralihan.

- (i) $\{w \in \{a, b\}^* : w \text{ does not end in } ba\}$.

$\{w \in \{a, b\}^* : w \text{ tidak diakhiri dengan } ba\}$.

- (ii) $\{w \in \{a, b\}^* : \#_a(w) \equiv_3 0\}$.

$\{w \in \{a, b\}^* : \#_a(w) \equiv_3 0\}$.

- (iii) $\{w \in \{a, b\}^* : w \text{ has even length and contains the substring } ab\}$.

$\{w \in \{a, b\}^* : \text{panjang } w \text{ adalah genap dan mengandungi sub rentetan } ab\}$.

- (iv) $\{w \in \{0, 1\}^* : w \text{ corresponds to the binary encoding (without leading 0's) of natural numbers that are evenly divisible by 4}\}$.

$\{w \in \{0, 1\}^* : w \text{ sepadan dengan pengekodan binari (tanpa awalan 0) nombor asli yang boleh dibahagi sama rata oleh 4}\}$.

(8/100)

- (c) Construct a DFA using transition diagram for the following languages, L_3 and L_4 as indicated by the regular operator(s).
 L_1 is the language from 2(b)(i) and L_2 is the language from 2(b)(ii).

*Reka suatu DFA dengan menggunakan gambar rajah peralihan untuk bahasa-bahasa berikut, L_3 dan L_4 , mengikut operasi biasa yang dinyatakan.
 L_1 merupakan bahasa dari 2(b)(i) and L_2 merupakan bahasa dari 2(b)(ii).*

(i) $L_3 = \neg L_1.$

(ii) $L_4 = (\neg L_1) \cap L_2.$

(6/100)

3. (a) Write regular expressions for the following languages.

Tuliskan ungkapan-ungkapan malar untuk bahasa-bahasa berikut.

- (i) The set of binary string that has at least 3 characters, and the third character is 0.

Set rentetan binari yang mempunyai sekurang-kurangnya 3 aksara, dan aksara yang ketiga ialah 0.

- (ii) The set of binary string that contains at least three "1"s.

Set rentetan binari yang mengandungi sekurang-kurangnya tiga "1".

- (iii) The set of binary string that starts and ends with the same character.

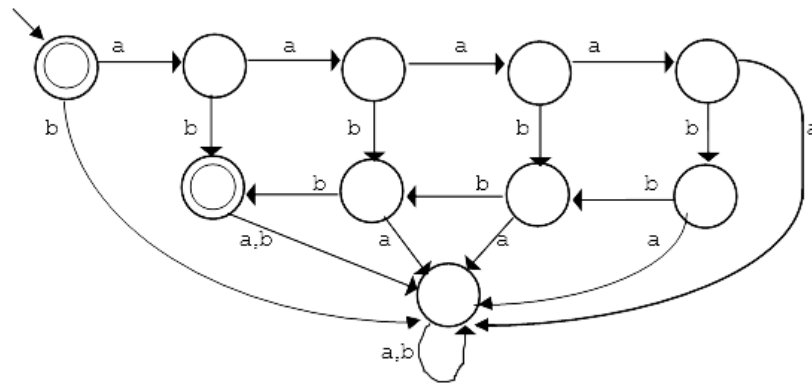
Set rentetan binari yang bermula dan berakhir dengan watak yang sama.

(6/100)

- (b) Convert the following DFA to regular expression. Describe the regular expressions using formal description.
 (Example of formal description: $\{w \in \{a, b\}^* : w \text{ contains at least one occurrence of the string } aa\}.$)

Ubah DFA berikut kepada ungkapan malar. Terangkan ungkapan malar menggunakan penerangan formal.

(Contoh penerangan formal: $\{w \in \{a, b\}^ : w \text{ contains at least one occurrence of the string } aa\}.$)*



(4/100)

4. Finite State Transducer (FST) is a special type of finite state machine.

Transducer Keadaan Terhingga (FST) adalah sejenis mesin keadaan terhingga khas.

- (a) Design a FST that converts an input "fat fat fat cat fat" into "thin thin thin cat thin".

Reka suatu FST yang menukarkan input "fat fat fat cat fat" kepada "thin thin thin cat thin".

(2/100)

- (b) Design a FST with input and output alphabet both $\{0,1\}$, so that the machine swaps each 1 in an even position to a 0. For example, for input 001101, it would output 001000.

Reka suatu FST dengan input and output berupa aksara $\{0,1\}$, yang dapat menukar setiap 1 yang berada di posisi genap kepada 0. Contohnya, input 001101, akan ditukar kepada output 001000.

(4/100)

5. Let $L = \{w \mid w \text{ is a palindromes over } \{a,b\}\}$. Use the pumping lemma to prove that L is regular language or otherwise.

Biar $L = \{w \mid w \text{ merupakan palindromes ke atas } \{a,b\}\}$. Guna "pumping lemma" untuk membuktikan bahawa L ialah bahasa biasa atau sebaliknya.

(4/100)

6. (a) Consider the following Context-Free Grammar (CFG):

Pertimbangkan Nahu Bebas-Konteks (CFG) berikut:

$$\begin{aligned} S &\rightarrow SpA \mid A \\ A &\rightarrow BmA \mid B \\ B &\rightarrow a \mid b \mid c \mid lSr \end{aligned}$$

S, A, and B are nonterminals, a, b, c, l, m, p, and r are terminals, and S is the start symbol. Draw a derivation tree according to the given grammar for the word **amlapbpcrmab**.

*S, A, dan B bukan terminal, a, b, c, l, m, p, dan r terminal, dan S ialah simbol mula. Lukiskan pohon terbitan berdasarkan nahu ini bagi perkataan **amlapbpcrmab**.*

(6/100)

- (b) Explain what it means for a Context-Free Grammar (CFG) to be ambiguous.

Terangkan apa yang dimaksudkan suatu Nahu Bebas-Konteks (CFG) menjadi taksa.

(4/100)

- (c) Is the following CFG ambiguous?

Adakah CFG yang berikut taksa?

$$\begin{aligned} A &\rightarrow AaA \mid AbA \mid B \\ B &\rightarrow c \end{aligned}$$

where A and B are nonterminals, A is the start symbol, a, b, and c are terminals.

di mana A dan B bukan terminal, A ialah simbol mula, a, b, dan c ialah terminal.

If yes, show two different leftmost derivations for the word *cacac*. If no, explain why.

Jika ya, tunjukkan dua terbitan paling kiri bagi perkataan cacac. Jika tidak, terangkan kenapa.

(6/100)

- (d) Describe the steps a Turing machine would use to recognize each of the following languages.

Huraikan algoritma suatu Mesin Turing yang akan digunakan untuk mengenali setiap daripada bahasa-bahasa berikut.

- (i) $L = \{b_n \# b_m \# b_{n+m} \mid b_i \text{ is } i \text{ in binary, } i \geq 1\}$. Example strings in this language include $0\#1\#1$, $11\#100\#111$, and $101\#110\#1011$.

$L = \{b_n \# b_m \# b_{n+m} \mid b_i \text{ ialah } i \text{ dalam binari, } i \geq 1\}$. Contoh rantaian-rantaian dalam bahasa ini termasuk $0\#1\#1$, $11\#100\#111$, dan $101\#110\#1011$.

(6/100)

- (ii) $L = \{w \mid w \text{ is a string of the form } (0^*1)^*, \text{ with all sets of 0's the same length}\}$. Examples strings of this language are ϵ , 0101 , 001001 , 00010001 , ...

$L = \{w \mid w \text{ ialah rantaian bagi bentuk } (0^*1)^*, \text{ dengan semua set bagi 0 sama panjang}\}$. Rantaian-rantaian bagi bahasa ini ialah ϵ , 0101 , 001001 , 00010001 , ...

(4/100)

- (iii) $L = \{a^i b^j c^k d^l \mid i = k \wedge j = l, i, j, k, l \geq 1\}$. Examples of string in L include $aabccd$, $abbbccddd$, $aaaabbbbccccddddd$, etc.

$L = \{a^i b^j c^k d^l \mid i = k \wedge j = l, i, j, k, l \geq 1\}$. Contoh-contoh rantaian dalam L $aabccd$, $abbbccddd$, $aaaabbbbccccddddd$, dll.

(4/100)

7. (a) The following is the formal description of the Pushdown Automata (PDA) that recognizes the language $\{0^n 1^n \mid n \geq 0\}$. Let M_1 is a PDA be $(Q, \Sigma, \Gamma, \delta, q_1, F)$, where

$$\begin{aligned} Q &= \{q_1, q_2, q_3, q_4\}, \\ \Sigma &= \{0, 1\}, \\ \Gamma &= \{0, \$\}, \\ F &= \{q_1, q_4\}, \text{ and} \end{aligned}$$

Berikut ialah huraian formal bagi Automata Tolakbawah (PDA) yang mengenali bahasa $\{0^n 1^n \mid n \geq 0\}$. Katakan M_1 ialah $(Q, \Sigma, \Gamma, \delta, q_1, F)$, di mana

$$\begin{aligned} Q &= \{q_1, q_2, q_3, q_4\}, \\ \Sigma &= \{0, 1\}, \\ \Gamma &= \{0, \$\}, \\ F &= \{q_1, q_4\}, \text{ dan} \end{aligned}$$

δ is given by the following table, wherein blank entries signify \emptyset .

δ ditunjukkan dalam jadual berikut, di mana masukan kosong menandakan \emptyset .

Input:	0			1			ε		
Stack:	0	\$	ε	0	\$	ε	0	\$	ε
q ₁									{(q ₂ , \$)}
q ₂	{(q ₂ , 0)}			{(q ₃ , ε)}					
q ₃				{(q ₃ , ε)}			{(q ₄ , ε)}		
q ₄									

Draw a state diagram of PDA for the above language.

Lakarkan satu rajah keadaan PDA bagi bahasa di atas.

(8/100)

(b) Consider the following Pushdown Automaton (PDA) P:

Pertimbangkan Automata Tolakbawah (PDA) P berikut:

$$P = (Q = \{q_0, q_1, q_2\}, \Sigma = \{a, b, c\}, \Gamma = \{a, \#\}, \delta, q_0, Z_0 = \#, F = \{q_2\})$$

where the transition function δ is given by

di mana fungsi transisi δ diberikan dengan

$$\begin{aligned} \delta(q_0, a, \#) &= \{(q_0, a\#)\} \\ \delta(q_0, c, \#) &= \{(q_0, \#)\} \\ \delta(q_0, a, a) &= \{(q_0, aa)\} \\ \delta(q_0, b, a) &= \{(q_1, q)\} \\ \delta(q_0, c, a) &= \{(q_0, a)\} \\ \delta(q_1, c, \#) &= \{(q_1, \#)\} \\ \delta(q_1, b, a) &= \{(q_1, q)\} \\ \delta(q_1, c, a) &= \{(q_1, a)\} \\ \delta(q_1, q, \#) &= \{(q_2, \#)\} \\ \delta(q, w, z) &= \emptyset \text{ everywhere else / mana-mana sahaja} \end{aligned}$$

Acceptance is by final state.

Penerimaan ialah dengan keadaan akhir.

(i) Which of the following words are accepted by the PDA P?

Yang manakah daripada perkataan berikut diterima oleh PDA P?

- acabbc
- abcabc

(2/100)

- (ii) For those words that are accepted by the PDA P , provide a sequence of Instantaneous Descriptions (IDs) leading to an accepting configuration as evidence. For those words that are not accepted, explain why there is no sequence of IDs leading to an accepting configuration.

Bagi perkataan yang diterima oleh PDA P , berikan suatu urutan bagi Penerangan Segera (IDs) yang membawa kepada suatu konfigurasi yang diterima sebagai bukti. Bagi perkataan yang tidak diterima, terangkan mengapa tiada urutan IDs yang membawa kepada suatu konfigurasi yang diterima.

(10/100)